DOCUMENT RESUME

ED 190 631 TM 800 426

AUTHOR Willson, Victor L.; Putnam, Richard R.
TITLE A Meta-analysis of Pretest Sensitization Effects in Experimental Design.
PUB DATE [79]
NOTE 23p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Gains; *Attitude Change; Control Groups; Experimental Groups; *Pretests Posttests; *Research Design; *Statistical Bias; Time Factors (Learning)
IDENTIFIERS *Effect Size; *Meta Analysis; Random Effects

ABSTRACT

A meta-analysis of 29 separate studies investigating pretest effects was conducted. Outcomes of the studies (achievement gains or attitude improvements) were computed as standardized differences between pretested and non-pretested groups. Eleven other variables were coded for each outcome. Initial descriptive statistics were indicative of differences between randomized and nonrandomized studies, so all further analyses were based on the 110 randomized group outcomes. For all outcomes, the average effect size was .34, indicating the general elevating effect of pretest on posttest. Cognitive outcomes were raised .51, attitude outcomes .33, and others about .27 standard deviations. Seventy percent of all effects were positive, and 93% of the cognitive effects were positive. Duration of time between pre- and post-testing was not related to effect size, although effect size is generally smaller for durations over one month. Year of publication, sample size, presence of experimental treatment, and sameness or difference of pretest and posttest were not significantly related to effect size. Researchers must continue to include pretest as a design variable when it is present, and to estimate its effect. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED190631

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A Meta-analysis of Pretest Sensitization Effects in Experimental Design

Victor L. Willson Richard R. Putnam

Texas A & M University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

V. Willson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 800426

Running head: A Meta-analysis of Pretest

Abstract

A meta-analysis of outcomes from studies investigating pretest effects was conducted. All outcomes were computed as standardized differences between pretested and non-pretested groups. Eleven other variables were coded for each outcome. Initial descriptive statistics were indicative of differences between randomized and nonrandomized studies, so all further analyses were based on randomized group outcomes (n=110). For all outcomes the average effect size was +.34, indicating the general elevating effect of pretest on posttest. Cognitive outcomes were raised .51, attitude outcomes .33, and others about .27 standard deviations. Seventy percent of all effects were positive, and 93% of the cognitive effects were positive. Duration of time between pre- and post-testing was not related to effect size, although effect size is generally smaller for durations over one month. Year of publication, sample size presence of experimental treatment, and sameness or difference of pretest and posttest were not significantly related to effect size. Researchers must continue to include pretest as a design variable when it is present, and to estimate its effect.

A Meta-analysis of Pretest Sensitization Effects in Experimental Design

Introduction

The pretest is a common condition of experimental and nonexperimental research. Campbell and Stanly (1963) discussed it in their landmark review of experimental design considerations Bracht and Glass (1968) and Welch and Walberg (1970) have presented more recent reviews of the effect of pretesting on posttest scores. Jaeger (1975) has advocated the use of pretests for many evaluation and research applications. The independent experimental contribution of pretests to dependent variables is unknown. Welch and Walberg's (1970) summary listed previous results in the usual review format of significance or non-significance. They concluded the long-term cognitive effects are small or nil, while there may be short term effects. These effects, they suggested, are greater for attitude tests than for cognitive tests.

The recent work of Glass (1976) and his colleagues and students in statistical review of research, termed meta-analysis, allows a more systematic analysis of the pretest experimental effect. Instead of relying on binary decisions about significance, the pretest effect is examined as a standardized difference between pretested and non-pretested groups. The purpose of the study is to estimate the pretest effect utilizing meta-analytic techniques and to investigate moderator variables associated with the potential pretest effect.

Procedures

Literature Search

In order to explore pretest sensitization and to determine the nature of its effect the Meta-Analysis techniques for integrating existing research findings were utilized as proposed by Glass (1977). Briefly summarized, this approach requires that either all studies or a sufficiently large representative sample on a given topic be located. Each study is treated as an individualized unit of analysis with multiple dependent variables and/or multiple comparison groups considered as separate units.

The method utilized in compiling the literature for examination was two-fold. The Automated Information-Retrieval Service (AIRS) at Texas A&M was employed to retrieve relevant information from the Educational Resources Information Center (ERIC), Psychological Abstracts, and the Social Science Citation Index (SOCIAL SCISEARCH). Secondly, references cited in the Welch and Walberg (1970) article were reviewed and a branching technique was utilized in an effort to locate all possible literature of interest to the present study.

When the task of locating the material for study was complete, the research reports were screened for appropriateness of content, and articles focusing on outcomes other than human educational and psychological phenomena were omitted (e.g.

animal studies. social interaction studies, neonatal response studies). Also, studies in which all groups were pretested were omitted.

Variables Coded

The variables extracted from each study in the literature search were year of publication, subjects' grade in school or age, sex of subjects (all males, all females, or mixed), duration of time betweeen pretest and post-test, citation (or not) of test reliability, sample size, randomization in selection of subjects (or not), random assignment of subjects to groups (or not), presence or absence of relevant treatment between pretest and posttest, category of dependent variable, similarity or dissimilarity of pretest and posttest, and effect size of pretest on posttest. The variables were chosen on the basis of theoretical relevance to the pretest effect, interest for experimental design, and utility in previous meta-analysis. For example, year of publication has been shown by Smith and Glass (1977) and Glass, Smith, and Barton (Note 1) to be related to effect size in several areas of research. Smith and Glass (1977) showed a similar effect for duration of treatment on psychoterapy outcomes. They also compared randomized with non-randomized experiments, demonstrating no difference in psychotherapy outcome effect sizes.

Dependent variables in each study were first coded by name

of test or scale, but it soon became apparent that a larger categorization was needed. Most outcomes were easily categorized as cognitive (achievement test score) or attitudinal. While four other categories were utimately distinguishable (physical performance, rating of performance, projective score, and choice preference). The number of outcomes for these was found to be small compared with those of the first two categories.

Duration of time between pre-and post testing was categorized ex post facto into four categories: hours, with duration of two hours or less; days, with duration less than seven; weeks, with duration less than two, and months, with duration less than or equal to nine. No studies researched in this paper investigated pretest effects for duration greater than nine months.

Effect sizes were computed from summary statistics reported by authors. Group means were required in all cases except when t-statistics (or F-statistics with one degree of freedom) were reported. Group sample sizes were also necessary. Each experimental design was reduced to one or more of the following forms (see Campbell and Stanley, 1963):

I.	R	O X O	II.	R	O O
	R	X O		R	O
III.		O X O	IV.		O O
		X O			O

Thus if no comparison of group means conformed to type I-IV, the study was ignored. For each form I-IV the effect size was defined as

$$\text{effect size} = \frac{\bar{X}_{\text{pretested}} - \bar{X}_{\text{pretested}}}{S_{\text{not pretested group}}} \quad , (1)$$

When the standard deviation of the not-pretested group was unavailable and not recoverable, a pooled variance statistic was used. If mean square error in a study was available, its square root served as standard deviation. If a t-statistic or F-statistic was reported the following transformation allowed estimation of the effect size:

$$\text{Effect} = t\,(1/n_1 + 1/n_2)^{\frac{1}{2}} \quad , \quad (2)$$

$$F = t,$$

where n_1, n_2 = sample sizes for the two groups.

It is possible some bias results in such cases if the the pretest affects variance of posttest scores. A test of this hypothesis was made by forming an F-ratio of pretested group variance divided by unpretested group variance for twenty-three identifiable cases. Two of the twenty-three (8.7%) were significant at $p = .05$, indicating no variance bias. In some designs it was possible to estimate mean square error from the means of two or more control groups, none of which was pretested (all were required to be of the same form for estimation). Mean square between these groups was computed and the square root taken to be

an estimate of group variability in the absense of pretest effect. The effect was then computed using equation (2). Irrelevant crossing variables were eliminated in factorial designs by averaging across them to obtain marginal means appropriate to forms I to IV.

In some designs several effects were computable utilizing different groups. This was commonly the result of several different treatments being administered. Each possible effect was coded in these designs. When several dependent variables were examined in a study a separate effect was computed for each dependent variable. This resulted in a few studies yielding a large number of effects. The number of effects per study varied from one to twenty-three. In some studies dependent variables were all of one category (eg. attitude) which in others dependent variable represented two or three categories of outcome.

The sign of the effect size is of great importance to interpretation of results. For cognitive outcomes a positive effect implies a higher test score for the pretested group. For attitude outcomes positive effect was defined as a more favorable attitude toward the subject. Sign of the effects was reversed to fit this definition. For behaviors the positive direction reflected the experimental aim of the study; for example avoidance behaviors should decrease in a systematic desensitization for fear of snakes. Thus, fewer avoidance behaviors was a positive outcome in this case. In a curriculum study the

the number of activities of high school students related to science was counted. An increase in this number was taken as a positive outcome.

Data Analysis

The first analysis was the extraction of descriptive statistics. Of special importance are the distributions of effects for randomized and non-randomized studies. If the distributions are statistically nondistinguishable then they may be pooled for subsequent analyses, as Smith and Glass (1977) did for psychotherapy outcomes. If different, the nonrandomized effects are presumed to have an unknown bias and are not further investigated.

In either case further analysis is used to examine the relationship of each coded variable to effect size, singly and multiply. Correlation, multiple regression and analysis of variance are all applied, depending on the nominal or interval character of the variables.

Results

Twenty-nine separate studies were analyzed which all had as a primary focus examination of the pretest effect. From these studies 140 separate effects were extracted and coded with the other independent variables. Summary statistics for these effects are given in Table 1, along with summary statistics for

subsets of the studies based on randomized or nonrandomized assignments and to pretest conditions. The nonrandomized effects have a different mean ($t = 2.44$, $df = 81.8$, $p<.02$, Satterthwaite's approximation using unequal variances) and different variance ($F = 3.03$; $df = 1.09.29$, $p<.01$) from the nonrandomized effects. Consequently, it is concluded that nonrandomized study effects are biased systematically and are excluded from further analysis.

Insert Table 1 about here

Summary statistics for randomized study effects are given in Table 2. Effects of subjects who received a treatment (Form I) had a mean of .31, not significantly different from untreated subject effects, mean .39. Both distributions are positively skewed.

A subset of the pretest sensitization literature has been concerned with the interaction of pretest condition and treatment. Globally, this hypothesized interaction does not appear to be large if it exists at all, since the effect size for treated and untreated groups are so similar.

The time between pretest and posttest as measured in hours, days, weeks and months is not linearly related to effect size. The pattern of means suggests small effects for delays of less than a day or over one months, but there are many gaps in the

intervals observed in the studies. Most short delays observed were of duration less than one hour. All studies with delays over a few hours and less than week utilized two or three day durations. Other studies used delays of seven, ten, and eleven and twelve days, while longer term studies used two week, thirty-five days, forty-five days, ten weeks, ninety-one days, and nine months. A systematic investigation of delay time between pretest and posttest would help to clarify the effect size.

Category of outcome yields some slightly surprising results. Cognitive effects under pretest conditions are over one half standard deviation above nonpretested groups, and 93% of the effects are positive. For attitude outcomes the effect mean is about .33 with 62% of the effects positive. This is a reversal of the conclusion of Bracht and Glass (1968) and Welch and Walberg (1970). Closer examination of the cognitive outcomes provides some useful results, given in Table 3. No clear conclusion about delay between pretest and posttest is possible, given the large effects which occurred for duration of one week's order. Pretest effect appears to be larger for pretest similar to posttest, and pretest effect is much larger for untreated groups than for treated groups.

For attitude outcomes many conclusions made about cognitive outcomes are reversed. Pretest effect appears larger for delays on the order of two days. Pretest effect is greater on different content posttests, and there is little difference in effect size between

treated and untreated groups.

Sex of subjects was not investigable since the overwhelming majority of studies (all but two) had mixed sex groups. Grade or age level of subjects is significantly related to effect ($F=5.09$, $df = 6, 103$; $p < .05$) but the significance is due to two cognitive outcomes observed with preschool children. In the absence of these effects overall, grade differences and differences within cognitive and attitude categories are all nonsignificant.

Year of publication, sample size, and design characteristics (report of reliability, random selection of subjects from population) were all nonsignificantly correlated with effect size. No multiple regressions were statistically or practically (R-square = .05) significant.

Discussion

The results of this meta-analysis of pretest sensitization effects provide two conclusions which have implications for educational, psychological and sociological research. The first is that there is a general pretest effect which cannot be safely ignored. Nonrandomized studies with pretests must be viewed with additional suspicion, since the data reported here are indicative of systematic bias due to pretest in such studies.

The second conclusion about pretest effects is that they appear not to be uniform across psychological domains. The studies reported here are not sufficiently exhaustive to provide definitive statements

about conditions for variation of pretest sensitization, but promising areas of investigation are change in cognitive effects over grade and age levels, transfer of pretest effect to different posttests, and change in effect over delay time between pretest and posttest from hours to months.

Reference Note

Glass, G. V., Smith, M. L. and Barton, M. A. Methods of integrative analysis (Report on Grant #NIE-G-78-0148) Boulder: University of Colorado, Laboratory of Educational Research, 30 August, 1979.

References

Anderson, E. J. The effects of combined learning cues on teacher learning outcomes. Journal of Research in Science Teaching, 1978, 15, 221-226.

Bracht, G. H. and Glass, G. V. The external validity of experiments. Am. Ed. Research Journal, 1968, 5, 437-474.

Campbell, D. T. and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand-McNally, 1966.

Cowan, G. and Komorita, S. S. The effects of forewarning and pretesting on attitude change. Educational and Psychological Measurement, 1971, 31, 431-439.

Entwisle, D. R. Interactive effects of pretesting. Educational and Psychological Measurement, 1961, 21, 607-620.

Gaskell, G. D., Thomas, E. A. C., and Farr, R. M. Effects of pretesting on measures of individual risk preference. Journal of Personality and Social Psychology, 1973, 25, 192-198.

Greitzer, G. & Jeffrey, W. E. Negative effects of pretest in training conservation of length. Developmental Psychology, 1973, 9, 435.

Hartley, J. The effect of pretesting on posttest performance. Instructional Science, 1973, 2, 193-214.

Hunt, R. G. & Lichtman, C. M. Pretest influences in evaluating organizational effects of a supervisory counseling training program. Journal of Applied Behavorial Science, 1972, 8, 503-507.

Israel, A. C., Becker, R. E., & Neilans, T. H. Contribution of pretesting to several measures of semantic desensitization effectiveness. Journal of Consulting and Clinical Psychology, 1977, 45, 1197-1198.

Jaeger, R. M. Some new developments and discoveries for evaluative analysis. CEDR Quarterly, 1975, 9, 3, 3-7.

Jones, J. M. Reactive effects of pretesting and test length in attitude research. Psychological Reports, 1973, 33, 107-114.

Kazdin, A. E. Effect of suggestion and pretesting on avoidance reduction in fearful subjects. Journal of Behavior Therapy and Experimental Psychiatry, 1973, 4, 213-221.

Kroeker, L. L. Pretesting as a confounding variable in evaluating an encounter group. Journal of Counseling Psychology, 1974, 21, 548-552.

Lana, R. E. Inhibitory effects of a pretest on opinion change. Educational and Psychological Measurement, 1966, 26 (1), 139-150.

Lana, R. E. The influence of the pretest on order effects on persuasive communications. Journal of Abnormal and Social Psychology, 1964, 69 (3), 337-341.

Lana, R. E. A further investigation of the pretest treatment effect. Journal of Applied Psychology, 1959, 43, 421-422.

Lana, R. E. Pretest treatment interaction effects in attitudinal studies. Psychological Bulletin, 1959, 56, 293-300.

Lana, R. E., & King, D. J. Learning factors as determiners of pretest sensitization. Journal of Applied Psychology, 1960, 44, 189-191.

Lana, R. E. & Menapace, R. H. Subject commitment and demand characteristics in attitude change. Journal of Personality and Social Psychology, 1971, 20, 136-140.

McCauley, C., Teger, A. I., and Kogan, N. Effect of the pretest in the risky shift paradigm. Journal of Personality and Social Psychology, 1971, 20, 379-381.

Nosanchuk, T. A. & Hare, R. D. Word recognition threshhold as a function of pretest sensitization. Psychonomic Science, 1966, 6, 51-52.

Pauling, F. J. & Lana, R. E. The effects of pretest commitment and information upon opinion change. Educational and Psychological Measurement, 1969, 29, 653-663.

Rosnow, R. L., Holper, H. M., & Gitter, A. G. More on reactive effects of pretesting in attitude research: Demand characteristics or subject commitment? Educational and Psychological Measurement, 1973, 33, 7-17.

Rosnow, R. L. & Suls, J. M. Reactive effects of pretesting in attitude research. Journal of Personality and Social Psychology, 1970, 15, 338-343.

Shearn, D. F. & Randolph, D. L. Effects of reality therapy methods applied in the classroom. Psychology in the Schools, 1978, 15, 79-82.

Smith, M. L. and Glass, G. V. Meta-analysis of psychotherapy outcome studies. American Psychologist, 1977, 32, 752-760.

Solomon, R. L. An extension of control group design. Psychological Bulletin, 1949, 46, 137-150.

Spanos, N. P., McPeake, J. D. & Carter, W. Effects of pretesting on response to a visual hallucination suggestion in hypnotic subjects. Journal of Personality and Social Psychology, 1973, 28, 293-297.

Welch, W. W. & Walberg, H. J. Pretest and Sensitization effects in curriculum evaluation. American Education Research Journal, 1970, 7, 605-614.

Table 1

Summary Statistics for Pretest

Effect Size Distribution

	N	Mean	S.D.	Percentage Positive Outcomes	Max	P_{90}	P_{50}	P_{10}	Min
All Studies	140	.26	1.02	67	4.46	1.15	.09	-.50	-2.35
Non-randomized Studies	30	-.04	.63	57	1.15	.44	.04	-.61	-2.35
Randomized Studies	110	.34	1.09	69	4.64	1.6	.09	-.50	-2.30

Table 2

Summary Statistics of Randomized Study

Effects for Selected Independent Variables

	N	Mean	S.D.	Percentage Positive Outcome	Max	P_{90}	P_{50}	P_{10}	Min
Treated Groups	70	.31*	1.06	73%	4.16	1.60	.06	-.50	-2.30
Untreated Groups	40	.39*	1.14	65%	4.46	1.41	.15	-.51	-1.63

	N	Mean	Omnibus F-Statistic	Mean Square Error
Duration				
Hours	30	.25	1.57	1.12
Days	38	.44*		
Weeks	15	.81*		
Months	26	.11		
Outcome Category				
Cognitive	27	.51*	.62	1.19
Attitude	58	.33*		
Other	25	.17		
Pretest-Posttest				
Same	83	.29	.56	1.19
Different	27	.47*		

* P < .05, Ho: Mean = 0

	N	Mean	Omnibus F-Statistic	Mean Square Error
Grade/Age of Subjects			5.09	.969
Adult	8	.20		
College	74	.34		
High School	4	.18		
Grade 11	12	.04		
Grade 9	8	.12		
Grade 4	2	.10		
Pre-School	2	4.06		

Correlation with Effect	N	R
Year of Publication	110	.008
Sample Size	110	-.128
Reliability of Outcome reported? (1=yes 0=no	110	-.078
Random selection of Subjects from Population? (1=yes, 0=no)	110	-.095

Table 3

Summary statistics of Randomized Study Effects for Cognitive and Attitude Outcomes

	Cognitive Outcomes		Attitude Outcomes	
	N	Mean	N	Mean
Duration				
Hours	16	.28	9	.03
Days	2	.17	30	.58*
Weeks	2	4.06*	13	.31
Months	7	.10	6	-.13
Pretest-Posttest				
Same	20	.63*	41	.18
Different	7	.15	17	.68*
Treatment				
Treated Group	22	.39	32	.37
Untreated Group	5	1.01*	26	.27
Grade/Age of Subjects				
Adult	-	-	2	-.35
College	10	.39	52	.38
High School	4	.18	--	--
Grade 11	--	--	4	-.02
Grade 9	8	.12	--	--
Grade 4	2	.10	--	--
Pre-school	2	4.04*	--	--

* P < .05; Ho: Mean = 0